Plant Pathology (2018) Doi: 10.1111/ppa.12942

Distribution and phylodynamics of papaya ringspot virus on *Carica papaya* in Cuba

D. Cabrera Mederos[abc*†], F. Giolitti[b], C. Torres[cd] and O. Portal[e†]

[a]*Departamento de Agronomía, Facultad de Ciencias Agropecuarias, Universidad Central 'Marta Abreu' de Las Villas, Santa Clara, Cuba;* [b]*Instituto de Patología Vegetal (IPAVE), Centro de Investigaciones Agropecuarias (CIAP), Instituto Nacional de Tecnología Agropecuaria (INTA), Av. 11 de Septiembre 4755, X5020ICA Córdoba;* [c]*Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Godoy Cruz 2290, C1425FBQ CABA, Buenos Aires;* [d]*Cátedra de Virología, Facultad de Farmacia y Bioquímica, Universidad de Buenos Aires, Junin 956, 4 piso, Ciudad Autónoma de Buenos Aires, C1113AAD, Argentina; and* [e]*Departamento de Biología, Facultad de Ciencias Agropecuarias, Universidad Central 'Marta Abreu' de Las Villas, Santa Clara, Cuba*

Orchard and garden papaya crops grown in 47 Cuban municipalities were surveyed from 2008 to 2013, revealing the widespread distribution of papaya ringspot virus (PRSV) in Cuba. Phylodynamic analyses performed with the partial coat protein gene of all Cuban PRSV-P isolates (34 sequences) and 107 sequences of isolates from the American continent and the Caribbean islands showed a most recent common ancestor in 1942 (95% highest posterior density, HPD 95% = 1911–1967). The substitution rate was estimated to be 7.7×10^{-4} substitutions per site per year (HPD 95% = 4.6×10^{-4} to 1.1×10^{-3}), which is equivalent to those detected in other RNA viruses. Demographic reconstruction of PRSV showed that viral diversity increased in the 1985–1990 period, which coincides with the implementation of extensive production practices. Moreover in Cuba, viral dispersion occurred from Mexico and other unknown ancestral locations. The spatiotemporal diffusion analysis proposed Mexico as an ancestral area for the origin of diversification in the American continent and suggests new dispersion events between American and Caribbean isolates. The observed widespread distribution, clear geographic grouping of Cuban isolates, virus growth and genetic diversity provide strong evidence of the PRSV dispersion patterns, which has implications for the control strategies of PRSV.

Keywords: papaya, phylodynamic analysis, PRSV

Introduction

Papaya (*Carica papaya*) is broadly cultivated in tropical and subtropical regions (Yeh *et al.*, 2007). Among fruits grown in Cuba, papaya ranks fourth in economic importance, with Maradol Roja being the main cultivar. Cuba produces about 197 000 t a year in about 6000 ha; these numbers make Cuba one of the 15 largest papaya-producing countries in the world (FAO, 2017).

Papaya ringspot virus (genus *Potyvirus*, family *Potyviridae*; PRSV; Wylie *et al.*, 2017) is the most economically important virus affecting papaya worldwide and reduces the plantation cycle to less than 1 year (Tripathi *et al.*, 2008). The virus is classified into two serologically indistinguishable strains (types P and W). The papaya strain (PRSV-P) is capable of systemically infecting papaya and species belonging to the family Cucurbitaceae, whereas the watermelon strain (PRSV-W) only infects cucurbits (Mansilla *et al.*, 2012).

PRSV is naturally transmitted by different species of aphids in a nonpersistent manner (Tripathi *et al.*, 2008). The abundance of these insects and their host in Cuba promotes the rapid spread of PRSV in papaya plantations and makes management very difficult. Symptoms induced by PRSV in papaya are vein clearing, mosaic and chlorotic mottle in leaves, and ringspot in fruits. The development of these symptoms depends on factors such as temperature, cultivars and viral isolates (Purcifull *et al.*, 1984; Davis *et al.*, 2004). Disease severity is related to the genetic and biological diversity of PRSV (Fernández-Rodríguez *et al.*, 2008). Moreover, epidemiological studies focusing on PRSV distribution are crucial to disease management (Noa-Carrazana *et al.*, 2006).

Genetic variability of PRSV is related to the geographical origin of the isolates (Silva-Rosales *et al.*, 2000; Bateson *et al.*, 2002). Gibbs *et al.* (2008) provided evidence of the Asian origin of PRSV and proposed that the introduction of the virus to the Americas occurred about 300 years ago; a similar hypothesis was formulated by Olarte-Castillo *et al.* (2011). Moreover, compared with the PRSV isolates in America (Chin *et al.*, 2007; Zambrana-Echevarría *et al.*, 2016), genetic diversity of PRSV populations is highest in India (Jain *et al.*, 2004), which coincides with the origin of this virus (Bateson *et al.*, 2002). Several investigations propose the use of

*E-mail: dcabreramederos80@gmail.com
†These authors contributed equally.

transgenic plants for PRSV management, although it requires a high homology between the transgene and the challenged virus coat protein (CP) gene (Fermin *et al.*, 2010). In Cuba, few genetic studies of PRSV have been conducted if transgenic plant programmes are envisaged.

PRSV was reported in Cuba for the first time in 1946 by the Agronomic Experimental Station of Santiago de Las Vegas (Acuña & Zayas, 1946). Despite the importance of papaya in Cuba, there are no records of the distribution, severity and phylodynamic behaviour of PRSV isolates; this information is important to gain knowledge of the virus impact and establish disease management programmes. Thus, a survey was conducted in all Cuban provinces from 2008 to 2013 to determine for the first time the distribution and diversity of PRSV in plantations of papaya cv. Maradol Roja. The distribution and severity of PRSV in geographically distant producing regions were analysed with the aim of contributing to disease management. Furthermore, phylodynamic methods were used to infer the rates of movement of viral lineages between geographical locations and to reconstruct the geographical locations of ancestral lineages among American and Caribbean isolates.

Materials and methods

PRSV symptoms in field-cultivated papaya plants

To determine PRSV severity, an evaluation scale was developed based on symptom severity in Maradol Roja plants cultivated under field conditions. Evaluations were conducted in papaya plantations located in a disease endemic area (Santo Domingo, Villa Clara). Papaya plants were grown in a nursery and transplanted to the field in January 2008. Then, symptoms were evaluated in the field every 15 days, for up to 240 days post-transplantation. The evaluated plantation consisted of 1666 plants with a 4 × 1.5 m planting distance, covering 1 ha. A detailed description of PRSV symptoms was performed in approximately 10% of total papaya plants, which were selected using a random zigzag approach. The evaluation was repeated once in a new plantation under similar conditions, starting in August 2009.

In both field evaluations, leaves from five papaya plants showing different symptoms were collected. Crude sap extracts from collected leaf samples were mechanically inoculated onto healthy papaya plants to confirm the causal agent. Plants were dusted with silicon carbide 600 mesh on the third and fourth youngest leaves and gently rubbed with 1 g papaya leaf extract in a 1:5 (w/v) dilution, containing 0.01 M potassium phosphate buffer (pH 7.0) and 0.1% sodium sulphite. Plants were maintained in an insect-free greenhouse. The symptom evaluation and virus detection were performed according to Cabrera Mederos *et al.* (2017).

Distribution, prevalence and severity of PRSV

To determine the distribution, prevalence and severity of PRSV in Cuba, a survey of plants with symptoms and sample collections were conducted in Maradol Roja plantations grown in privately and government-owned lands. In addition, papaya plants from gardens were also assessed but not considered for determining incidence and disease index. Evaluations were conducted from November 2008 to May 2013 in 70 locations corresponding to 47 municipalities from the 15 Cuban provinces (Table S1).

To determine the distribution and prevalence of PRSV, the collected papaya leaves were mechanically inoculated as described above for observation of symptoms, and part of the samples were stored at −80 °C until use. Virus was initially detected from field-collected leaves and inoculated plants using ELISA, according to Cabrera Mederos *et al.* (2017). For the molecular detection of PRSV, total RNA extraction was performed with RNeasy Plant Mini kit (QIAGEN), according to the manufacturer's recommendations. The cDNA was synthesized from 1 μg total RNA using the High Capacity cDNA Reverse Transcription kit (Applied Biosystems). The coat protein (CP) cistron of the PRSV was amplified using the BoCP_Fwd (5′-TCCAAGAATGAAGCTGTGGACGCTGGTT-3′) and BoCP_Rev (5′-TYAGTTGCGCATACCCAGGAGAGAGT-3′) primers, according to Cabrera Mederos *et al.* (2016). PCR conditions included an initial denaturation cycle at 94 °C for 2 min; followed by 34 cycles of denaturation at 94 °C for 30 s, annealing at 58 °C for 45 s and an extension at 72 °C for 1 min; with a final extension at 72 °C for 10 min. PCR was performed in a MasterCycler personal PCR machine (Eppendorf) with TopTaq DNA polymerase (QIAGEN). The distribution of PRSV was determined by dividing the locations where the virus was detected by the total locations evaluated.

All the papaya fields and gardens where PRSV was detected from the total sampled locations were taken into account to determine the prevalence of the virus. Assessments to calculate PRSV severity (incidence and disease index) in papaya plantations were performed in 10% of the planted area. Incidence of PRSV (I) was calculated as follows:

$$\mathrm{I}(\%) = n/N \times 100$$

where n is the total number of plants with symptoms and N is the total number of evaluated plants (Cooke, 2006). The disease index (DI) was calculated using the described symptoms and the developed scale. The formula

$$\mathrm{DI}(\%) = \frac{\sum(a \times b)}{N \times K} \times 100$$

was used, where a is the total number of plants at each scale level, b is the corresponding scale level, N is the total number of evaluated plants and K is the maximum scale level ($K = 5$) (Townsend & Heuberger, 1943). To use this scale score in the evaluation of the PRSV in field-cultivated papaya plants, the observations were made in papaya plantations from different provinces, and the results obtained by Cabrera Mederos *et al.* (2017) were also considered.

Sequencing and phylogenetic analysis

The obtained RT-PCR amplicons of the expected size (*c.* 850 bp) were purified and then directly sequenced in both directions by Macrogen Inc. The 25 new partial CP gene sequences of PRSV obtained in this work were deposited in GenBank; details of the origins and the sequences used for comparison are shown in Table S2. Sequences obtained were assembled with SeqMan (DNASTAR) and PRSV isolates were aligned using muscle, in mega v. 7.0 software (Kumar *et al.*, 2016). After sequence trimming, an 822 bp CP gene fragment was used to perform the phylodynamic analyses. The saturation was tested

using the method of Xia *et al.* (2003) implemented in DAMBE v. 6 software (Xia, 2017). The analysis was carried out on all sites and on third codon positions. In both cases, the index of substitution saturation (Iss) was significantly lower than the critical Iss value (Iss.c) ($P < 0.0001$), indicating that the dataset does not present significant saturation and is adequate for phylogenetic reconstruction. The model of base substitution was estimated using the JMODELTEST v. 2.1.9 software (Darriba *et al.*, 2012), according to the Akaike information criterion. The substitution model selected was GTR + G4 + I. The phylogenetic tree was obtained using Bayesian inference in MRBAYES v. 3.2.6 software (Ronquist *et al.*, 2012) in the CIPRES Science Gateway server (Miller *et al.*, 2010). Two Markov chain Monte Carlo (MCMC) computations were run twice for three million states and sampled every 5000 generations. Convergence was assessed by effective sample size values higher than 200 using the TRACER v. 1.6 software (Rambaut *et al.*, 2014), and the first 10% of generations were discarded as burn-in.

Phylodynamic analysis

The PRSV nucleotide sequences obtained in this work and others available in GenBank were subjected to a Bayesian coalescent analysis. The temporal signal of datasets was explored with a root-to-tip analysis in TEMPEST v. 1.5 (Rambaut *et al.*, 2016), showing a positive correlation between genetic divergence and sampling time. In addition, a date-randomization test has been performed with 20 replicas, using the TIPDATINGBEAST R package (Rieux & Khatchikian, 2017). This analysis showed no overlapping in the marginal posterior probability distributions of the time to the most recent common ancestor (tMRCA) and the mean substitution rate estimated with the real data and the replica, indicating significant temporal structure of the dataset (Figs S1 & S2). Thus, the substitution rate, the tMRCAs and the spatial dynamics in real timescales for PRSV isolates were estimated. The analyses were carried out using an appropriate substitution model (see above). Different combinations of molecular clocks (strict and uncorrelated lognormal) and demographic models (constant, exponential, logistic, expansional and Bayesian skyline plot (BSP)) were compared through Bayes factor using the marginal likelihood values estimated by the path sampling (PS) and stepping stone (SS) methods (Baele *et al.*, 2012). According to these results (Table S3), the models used were: uncorrelated lognormal molecular clock model (Drummond *et al.*, 2006) and the BSP coalescent model (Drummond *et al.*, 2005) implemented in the BEAST v. 1.8.4 software package (Drummond *et al.*, 2012). The analyses were run using BEAGLE library (Ayres *et al.*, 2012) in the CIPRES Science Gateway server (Miller *et al.*, 2010). A spatial diffusion process was modelled on time-measured genealogies over discrete sampling locations (country of origin; Lemey *et al.*, 2009). A Bayesian stochastic search variable selection (BSSVS) procedure was applied to obtain the set of spatial diffusion rates that explain the spatiotemporal process and a Bayes factor test was used to evaluate the significance of the spatiotemporal linkage between locations, using the SPREADD3 software (Bielejec *et al.*, 2016). The analysis was run for 400 million states, sampled once every 400 000 states. Convergence was assessed by effective sample size (ESS) values higher than 200, using TRACER v. 1.6 software (Rambaut *et al.*, 2014), and 10% of the sampling was discarded as burn-in. Uncertainty in parameter estimates was evaluated in the 95% highest posterior density (HPD 95%) interval.

Results

PRSV symptoms in field-cultivated papaya plants

In the evaluations performed in the papaya plantations located in a disease endemic area, the earliest PRSV symptoms were vein clearing in young leaves. This was followed by a mild mosaic characterized by small clear green zones observed on the adaxial leaf surface. Later, patches with light and dark green zones were formed, which resulted in mottled leaves. Then, slight deformation and dark green swelling zones were noticed in affected leaves. Finally, the most advanced disease symptoms were the simultaneous occurrence of an intense mosaic and leaf reduction. Oily spots were observed at the stem top and on petioles. A variable number of concentric rings in the earliest fruits were observed (Fig. S3). PRSV symptoms in the evaluated papaya plantations were similar on both planting dates. The presence of PRSV in papaya samples with any of the symptoms was confirmed in inoculated papaya plants using RT-PCR (data not shown). From the PRSV symptoms described, an evaluation scale of 0–5 was developed (Table 1), which was used to assess the virus severity in field-cultivated Maradol Roja papaya in Cuba.

Distribution, prevalence and severity of PRSV

Leaves of surveyed and inoculated papaya test plants, which developed the typical PRSV symptoms, were analysed by ELISA. In addition, the presence of PRSV was confirmed by amplification of an 850 bp fragment, which was not observed in symptomless samples. PRSV was identified in 36 of 70 locations (51.4%); numerical

Table 1 Evaluation scale corresponding to the papaya ringspot virus symptom severity in papaya plants cv. Maradol Roja cultivated under field conditions in Cuba.

Score	Description
0	No symptoms
1	Symptom onset: Vein clearing in young leaves or small concentric rings in fruits
2	Non-generalized symptoms: Mild mosaic confined to a small zone of the leaf or fruit with concentric rings. Symptoms affecting up to 25% of the affected organ
3	Generalized symptoms: Patches formed by clear and intense green zones alternating on the adaxial leaf surface; fruits with concentric rings. Symptoms affecting up to 50% of the plant
4	Severe symptoms: Mottled leaves alternating with green zones. Slight leaf deformation, dark green swelling zones and fruits with concentric rings. Symptoms affecting 51–75% of the plant
5	Very severe symptoms: Intense mosaic produced by chlorosis covering the entire leaf (chlorotic leaf), leaf distortion, oily spots well defined on the top of the stem and petioles. Distinctive concentric rings on fruits. Symptoms affecting more than 75% of the plant

tags were assigned to map locations where the disease was identified (Table 2; Fig. S4). Of the 111 papaya plantations and gardens surveyed, 55 points were positive for PRSV, with an overall prevalence of 49.5%.

The I and DI values of PRSV ranged from 0 to 100%, depending on the location and date of sampling. The highest I and DI values (100%) were recorded in San Luis (Pinar del Río), San Antonio de los Baños (Artemisa), Santo Domingo (Villa Clara), Sancti Spíritus (Sancti Spíritus) and Las Tunas (Las Tunas). In Melena del Sur and Nueva Paz (Mayabeque), Corralillo (Villa Clara), Trinidad (Sancti Spíritus), Rafael Freyre (Holguín) and Bayamo (Granma), I and DI values were lower than 60%. In papaya plantations from the other 34 locations, including plantations from Cienfuegos, Ciego de Ávila, Santiago de Cuba and Guantánamo provinces, the disease was not identified during the evaluations (data not shown). However, in Cienfuegos (Cienfuegos) and Palma Soriano (Santiago de Cuba), PRSV presence was detected in papaya plants cultivated in gardens (Table 2).

Genetic distance and phylogenetic analysis

The pairwise analysis of the Cuban PRSV isolates ($n = 34$) showed nucleotide sequence identity values ranging from 92.4% to 99.7%, with the lowest values being detected between isolates from different regions. The highest variability was observed between isolates from the western-central region and isolates from the east region: Melena del Sur-1–Campechuela (92.4%; MF041944–MF041959), Melena del Sur-1–Niquero (92.4%; MF041958), Melena del Sur-1–CbGR1 (92.5%; KC748224) and Trinidad–CbGR2 (92.5%; MF041951–KC748225) isolates (Fig. S5). The pairwise analysis of the amino acid sequences showed values ranging from 90.5% to 99.6%. The Bayamo isolate showed the highest variability, even with isolates of nearby areas. A tendency to the highest nucleotide identity values was observed among the isolates from the western-central region of Cuba, in most cases the highest level of conservation was detected among isolates from nearby regions.

The 25 partial CP gene sequences of PRSV obtained in this work and other sequences from Cuba and the American continent available in GenBank were analysed and a phylogeny was generated using Bayesian inference. In the phylogenetic tree, seven subgroups were observed, showing geographical clustering. Subgroup 1 included sequences from Brazil, Argentina and Jamaica, subgroup 2 included sequences from Puerto Rico and Venezuela, subgroup 3 included other sequences from Venezuela and Colombia, and subgroup 7 included other sequences from Puerto Rico, Mexico and Colombia. Sequences from Cuba were distributed among subgroups 4, 5 and 6 together with sequences from Mexico, Puerto Rico and Florida (US), and were segregated according to geographical origin (Fig. 1). The nucleotide sequences of Cuban isolates collected from 2010 to 2012 in the eastern region were included in the monophyletic subgroup 4. In this subgroup two lineages were formed (sequences from 2010 to 2011, and sequences from 2010 to 2012). The Cuban sequences collected from 2004 to 2013 in western-central regions were included in subgroup 5, except for an isolate from the eastern region (MF041962), which was closely related to sequences from Mexico and Puerto Rico. The isolate from Guáimaro (MF041952), collected from the east of Cuba, clustered in subgroup 6 with a Florida (US) isolate and other isolates from Mexico. The glutamic acid and lysine (EK) repeat boxes (I–IV), previously reported in the CP amino terminal region of PRSV (Silva-Rosales *et al.*, 2000), were detected in all analysed isolates (data not shown). In the isolates from Melena del Sur-1 and Santo Domingo-1 (both isolates from the western-central regions), the third box lacked an EK repeat. In the analysed sequences, this pattern was only observed in the Mexican and Cuban isolates (Fig. S6), as well as in one isolate from Puerto Rico (AF196838).

Phylodynamic analysis

The phylodynamic patterns of PRSV from the American continent and the Caribbean islands were studied under a Bayesian approach that considers the phylogenetic dispersion in time and space. The maximum clade credibility tree (MCCT) from the BEAST analysis, using the uncorrelated lognormal relaxed-clock model, showed similar groups to those of the previously obtained phylogenetic tree. The analysis of the spatiotemporal diffusion process of PRSV-P isolates (American continent and the Caribbean islands) showed a most recent common ancestor (mean) in 1942 (HPD 95% = 1911–1967), with a substitution rate of 7.7×10^{-4} substitutions per site per year (HPD 95% = 4.6×10^{-4} to 1.1×10^{-3}), which is equivalent to those of plant viruses (Pagán & García-Arenal, 2018). The spatiotemporal diffusion analysis proposed Mexico as an ancestral area for the American continent and the Caribbean islands (posterior probability of the ancestral state (PPAS) = 0.91; Fig. 2). PRSV sequences from Jamaica and Argentina clustered in subgroup 1 together with sequences from Brazil, which seems to be the ancestor of sequences from these countries (PPAS = 1.0). The Cuban sequences of subgroup 5 (western-central regions) were monophyletic with a high posterior probability value, dating to 1968 (HPD 95% = 1949–1986) and could be the result of a dispersal event from Mexico (PPAS = 0.99). This subgroup was associated with the Mexican isolates collected in 2000 from San Luis Potosí and Veracruz (northern and southern Gulf Coastal Plain), according to physiographic regions of Mexico (Noa-Carrazana *et al.*, 2006). Moreover, the Cuban sequences from subgroup 4 might correspond to an independent dispersion, starting their diversification in 1975 (HPD 95% = 1956–1989), although it was not possible to estimate their origin with confidence (posterior probability < 0.5). One of the Cuban isolates from the eastern region (MF041952) clustered with an isolate from Florida (US), and derived from

Table 2 Incidence and severity of papaya ringspot virus in papaya cv. Maradol Roja plantations assessed in major papaya-producing regions in Cuba.

Province	Municipality	Location	Date	ELISA/ PCR[a]	Symp./ Assessed[b]	I (%)	DI (%)	Map label[c]
Pinar del Río	San Juan y Martínez	El Cafetal[d]	Feb. 2009	+/+	1/1	—	—	1
	San Luis	San Luis	Feb. 2009	+/+	150/150	100	100	2
Artemisa	San Antonio de los Baños	Finca El Recodo	Jan. 2012	+/+	166/166	100	100	3
La Habana	Boyeros	Wajay	Feb. 2011	+/+	82/100	82	74	4
Mayabeque	Bejucal	Arsenal[d]	Sep. 2010	+/+	1/1	—	—	5
	San José de las Lajas	Finca Las Papas	Sep. 2010	+/+	128/200	64.18	64	6
	Melena del Sur	Los Pinos	Apr. 2009	+/+	10/60	16.70	14.70	7
			Jan. 2012	+/+	48/166	28.92	26.75	
	Güines	Güines[d]	Feb. 2009	+/+	1/1	—	—	8
	Nueva Paz	Nueva Paz	Feb. 2009	+/+	10/60	16.70	12.70	9
Matanzas	Jagüey Grande	Jagüey Grande	Mar. 2009	+/+	34/60	56.70	56.70	10
			May 2013	+/+	100/100	100	94	
Cienfuegos	Cienfuegos	Paraíso[d]	May 2010	+/+	1/1	—	—	11
Villa Clara	Corralillo	Motembo	Dec. 2012	+/+	65/222	29.28	29.19	12
				+/+	106/222	47.75	47.75	
	Santo Domingo	Santo Domingo	Nov. 2008	+/+	333/333	100	98	13
			Jan. 2010	+/+	250/250	100	100	
			Apr. 2010	+/+	150/250	100	88.50	
			Jan. 2011	+/+	166/166	100	93.40	
			Nov. 2011	+/+	174/222	78.38	53.55	
	Cifuentes	Braulio Coroneaux	Feb. 2012	+/+	63/100	63	34.60	14
	Santa Clara	Valle del Yabú	Oct. 2008	+/+	191/200	95.50	95.50	15
			Aug. 2012	+/+	9/115	7.83	6.09	
		Maleza	Aug. 2011	+/+	35/333	10.50	5.90	16
		Base Aérea	Mar. 2012	+/+	2/80	2.50	1.50	17
				+/+	2/115	1.74	1.04	
		El Gigante	May 2010	+/+	123/200	61.50	61.10	18
		Finca El Progreso	Mar. 2013	+/+	36/80	45	34	19
	Camajuaní	Vueltas	Feb. 2011	+/+	216/222	97.30	94.40	20
	Placetas	Placetas[d]	Nov. 2010	+/+	1/1	—	—	21
Sancti Spíritus	Sancti Spíritus	La Bija	Feb. 2010	+/+	10/60	16.7	12.0	22
		Finca Las Delicias	Feb. 2012	+/+	166/166	100	100	23
				+/+	66/66	100	100	
				+/+	66/66	100	100	
	Trinidad	Banao Toma de Agua	Nov. 2012	+/+	19/222	8.56	6.40	24
				+/+	56/240	23.33	21.17	
	Taguasco	Taguasco[d]	Feb. 2013	+/+	1/1	—	—	25
Camagüey	Guáimaro	Casa Azul	Feb. 2012	+/+	161/166	96.99	94.46	26
				+/+	14/88	15.91	8.41	
Las Tunas	Las Tunas	Becerra	Jun. 2011	+/+	6/65	9.23	7.38	27
				+/+	100/100	100	100	
				+/+	192/200	96.00	94.80	
	Puerto Padre	Parada	Jun. 2011	+/+	21/128	16.41	10.16	28
		Maniabón	Jun. 2011	+/+	30/45	66.67	53.78	29
				+/+	101/148	68.24	59.32	
Holguín	Calixto García	Buenaventura[d]	Feb. 2010	+/+	1/1	—	—	30
	Rafael Freyre	La Caridad	Dec. 2010	+/+	59/320	18.40	17.25	31
			Oct. 2011	+/+	19/200	9.50	6.10	
Granma	Niquero	Niquero[d]	Nov. 2012	+/+	1/1	—	—	32
	Campechuela	Campechuela[d]	Nov. 2012	+/+	1/1	—	—	33
	Manzanillo	Manzanillo[d]	Nov. 2012	+/+	27/30	90.00	90.00	34
	Bayamo	Huerto Las Bayamesas	Mar. 2012	+/+	43/185	23.24	19.46	35
				+/+	13/222	5.86	5.23	
				+/+	23/160	14.38	12.63	
Santiago de Cuba	Palma Soriano	Palma Soriano[d]	Nov. 2011	+/+	1/1	—	—	36

I, incidence; DI, disease index; —, I and DI values not calculated.

[a]Leaves from five papaya plants were tested in each plantation.

[b]Number of plants with symptoms/number of plants assessed.

[c]Map references used in Figure S4.

[d]Papaya plants from gardens.

a different dispersion event dating to 1963 (HPD 95% = 1941–1988), possibly from Mexico (PPAS = 0.67), within subgroup 6 (Fig. 2). In this subgroup two lineages were formed with the Mexican isolates collected from 1998 to 2000 (northern and southern Gulf Coastal Plain, Neovolcanic axis), and other isolates collected from 2000 (Neovolcanic axis, Pacific Coastal Plain and southern Sierra Madre; Fig. 1).

Demographic reconstruction of PRSV showed that viral diversity increased between 1944 and 1951, after which the viral population remained almost constant until the 1985–1990 period, when it increased slightly, then remained constant until the present (Fig. 2). The spatiotemporal diffusion processes showed the gradual dispersion of PRSV between countries, starting in Mexico as the most probable ancestral location in the American continent and the Caribbean islands (Fig. 3). Brazil appears to play an important role in the dispersion of PRSV to Argentina. A recent transmission event was observed between Venezuela and one isolate from Colombia, as they clustered together in subgroup 3. In addition, Bayes factors to identify well-supported rates between locations have been estimated (Table S4). This analysis reinforced the estimation that virus dispersed mainly from Mexico to different locations, coinciding with that observed in the MCCT.

Discussion

Mosaic, leaf deformation, oily spots on petioles and stem, and concentric rings on fruits were reported in this work, which are typical symptoms of this viral disease. In addition, fruit deformation was detected due to the intensity of the PRSV symptoms. The results obtained agree with those reported by other authors, who described chlorosis at infection onset, followed by mottling and leaf lamina reduction (Purcifull *et al.*, 1984; Tripathi *et al.*, 2008). The use of a scale for PRSV evaluation under field conditions has been widely used in the comparison of the efficiency of management methods in commercial papaya plantations, the selection of tolerant genotypes, presumptive detection of the disease in the field, and determination of resistance of transgenic papaya breeding lines (Crane *et al.*, 1995; Davis *et al.*, 2004). However, some PRSV symptoms did not fully match those descriptions of the disease progress in Cuban papaya plantations. In Cuba, necrosis induction has not been detected in the veins, and plant death is not considered a result of PRSV infection. Here, a scale is proposed to determine the severity of PRSV symptoms in papaya cv. Maradol Roja.

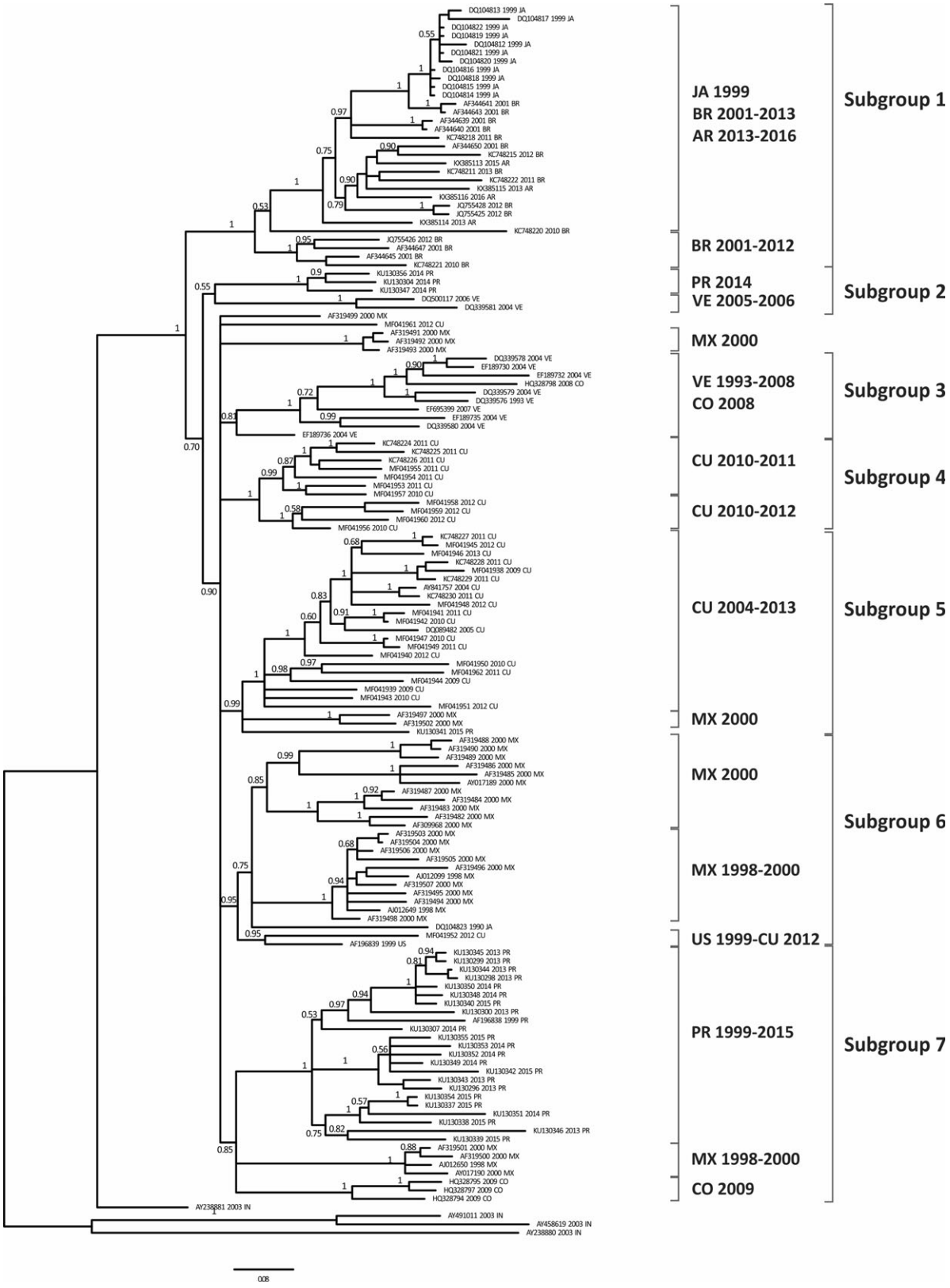
The high incidence and wide distribution of PRSV in Cuba could be the result of a high inoculum pressure and abundance of aphid vectors of this virus in most papaya-producing areas. Pinar del Río was one of the first Cuban provinces where papaya plantations were established during the 1930s (Roig, 1965). In San Luis (Pinar del Río), where high PRSV incidence and severity values were observed in a 300-day-old plantation, the virus was not detected in adjacent young plantations. This situation might be related to the occurrence of hurricanes Gustav (August 2008) and Ike (September 2008) through the province 6 months earlier. The strong winds and heavy rains caused by the hurricanes might have limited the disease incidence in the field because of its negative influence on aphid vector populations. The I values of PRSV in papaya plantations from the Citrus Company 'Jagüey Grande' (Matanzas) may be related to the presence of citrus as hosts of aphid vectors, which favours the spread of the disease (Holman, 1974). In Villa Clara province, the presence of insect vectors of PRSV in the evaluated areas, and infected papaya plantations in neighbouring areas, contribute to the spread of the virus. The cultivar Maradol Roja was obtained from the Instituto de Investigaciones de Viandas Tropicales in Santo Domingo, where its basic seed is produced, and the same areas are used for papaya plantations every year. This may be the reason for the high incidence and disease index (100%) in papaya plantations recorded in these areas.

The provinces located in the eastern sector of the country are characterized by high temperatures and the areas dedicated to papaya crop are smaller than western-central regions, mainly because the eastern sector is mountainous with limited availability of water for irrigation. Urban areas in Granma province (Niquero, Campechuela and Manzanillo), where PRSV infected plants were detected, are located near agricultural markets. These outlets where papaya fruit are traded are several miles away from papaya production areas. Accordingly, Lecoq *et al.* (2003) demonstrated the potential of cucurbit fruit as potyvirus transmission sources over long distances. This situation suggests the presence of infected fruits as the main route of PRSV transmission, because they are the inoculum for transmission by aphid vectors, which find host plants for reproduction in the gardens of these cities. Another dispersion route, which is more plausible, is the movement of infected papaya seedlings between regions. However, papaya plants infected with PRSV detected in these areas serve as a source of inoculum for virus dispersion.

Several studies have shown a clear correlation between the genetic differences among isolates and their

Figure 1 Majority rule consensus tree obtained from the Bayesian analysis of partial coat protein gene of PRSV isolates from the American continent and the Caribbean islands. Branch lengths are proportional to genetic distances and branch significance is indicated at nodes for posterior probability values higher than 0.5. Year of collection and country (abbreviated in uppercase letters) are indicated next to the accession number of the isolates. AR: Argentina, BR: Brazil, CU: Cuba, CO: Colombia, IN: India, JA: Jamaica, MX: Mexico, PR: Puerto Rico (US), US: Florida and VE: Venezuela. PRSV isolates from India were used as out-group.

geographical origin (Bateson *et al.*, 2002; Noa-Carrazana *et al.*, 2006; Yasaka *et al.*, 2017), which was observed among PRSV Cuban isolates. Most of the Cuban isolates were grouped into two subgroups (western-central and eastern regions). Rodríguez *et al.* (2014) analysed only seven available Cuban PRSV sequences and noted that



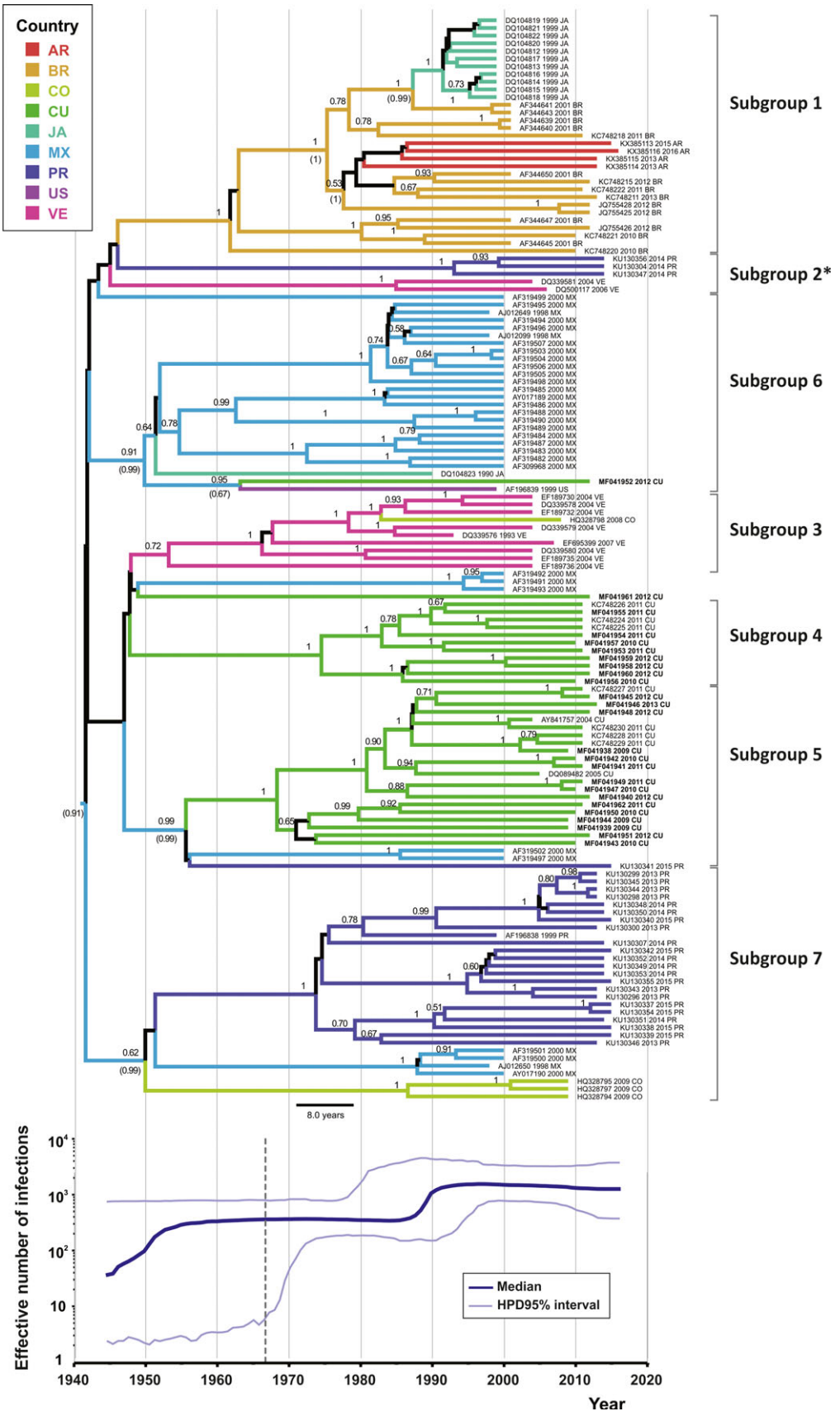

Country
AR
BR
CO
CU
JA
MX
PR
US
VE
Subgroup 1
Subgroup 2*
Subgroup 6
Subgroup 3
Subgroup 4
Subgroup 5
Subgroup 7
8.0 years
Effective number of infections
Median
HPD95% interval
Year

Figure 2 Maximum clade credibility tree and phylodynamic reconstruction for PRSV isolates from the American continent and the Caribbean islands. Branches are colour-coded online to represent the estimated geographical origin of analysed sequences, with black branches used to show groups with posterior values lower than 0.5. The dated tree is in the same temporal scale as the phylodynamic reconstruction shown beneath. Posterior values and probability of ancestral state (in parentheses) for relevant groups are shown at nodes. Year of isolation and country (abbreviated in uppercase letters) are indicated after the accession number of the isolates. AR: Argentina, BR: Brazil, CU: Cuba, CO: Colombia, JA: Jamaica, MX: Mexico, PR: Puerto Rico (US), US: Florida and VE: Venezuela. Sequences reported in this work are shown in bold. Asterisks indicate a group that was monophyletic in the Bayesian phylogenetic tree.

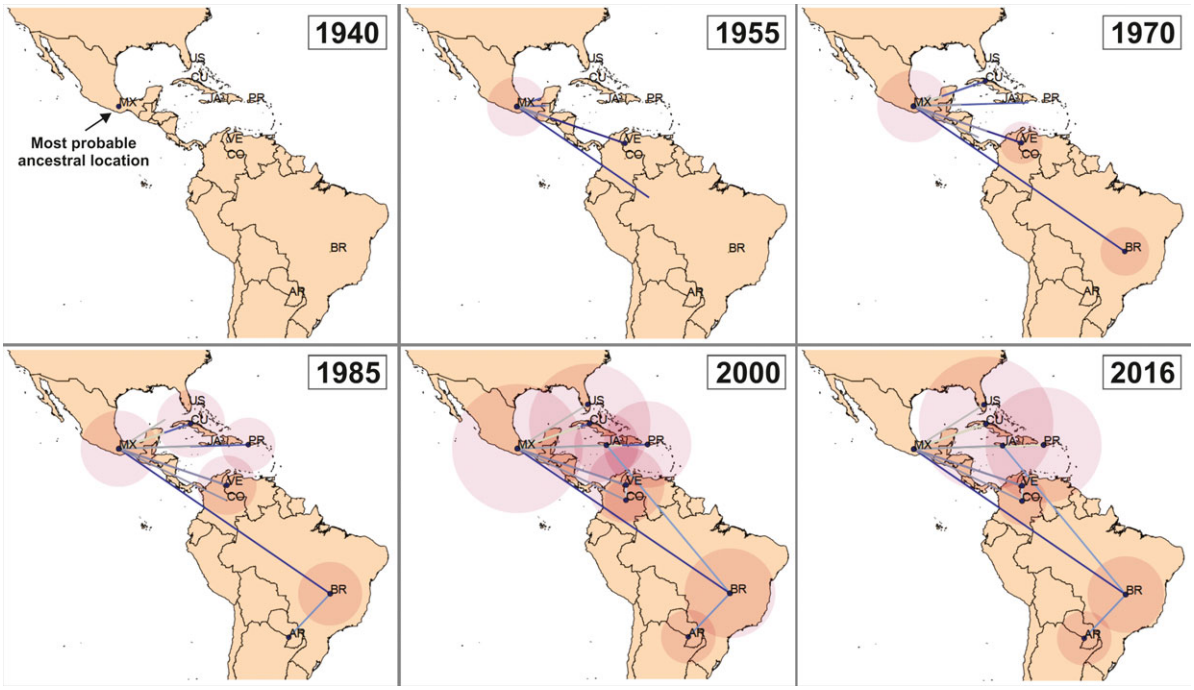


Figure 3 Spatiotemporal dispersion processes for PRSV in the American continent and the Caribbean islands. Lines represent the maximum clade credibility tree visualized on the maps. The dark–light gradient of line colours represents the relative age of the dispersal pattern (older to more recent, respectively). The large circles represent the relative lineage diversity in each period.

genetic conservation is greater among isolates from nearby regions. However, the results obtained in this study, through the analysis of a greater number of Cuban isolates, reflect cases in which this hypothesis is not supported. The Palma Soriano (MF041962) isolate, collected in the eastern region, showed high similarity with isolates from the western-central region. In Mexico, Noa-Carrazana *et al.* (2006) noted that the similarity between isolates collected in different regions was associated with the movement of genetic material. This fact is plausible, considering that in the western-central regions wide areas of papaya plantations were developed between 1930 and 1940 (Roig, 1965), which could have favoured the movement and introduction of genetic material. The grouping of isolates within a country according to their geographical origin has been reported in Mexico, Cuba and Brazil (Noa-Carrazana *et al.*, 2006; Rodríguez *et al.*, 2014). Nevertheless, this scenario was not observed in Indian isolates (Jain *et al.*, 2004). Similarly to Indian isolates, Puerto Rican isolates grouped independently of their geographical origin, although a higher level of conservation was observed among the latter isolates (Zambrana-Echevarría *et al.*, 2016). In this work, genetic variation was detected between isolates collected from the same areas (Boyeros-1 and -2, Melena del Sur-1 and -2, Santo Domingo-1 and -2), on different collection dates (2–6 years), which coincided with observations of Chin *et al.* (2007). These authors analysed sequences collected from Jamaica and Venezuela in different periods and associated the variability with the movement of germplasm and the implemented management practices.

The EK conserved blocks next to the DAG triplet in the CP have been previously detected in PRSV isolates (Silva-Rosales *et al.*, 2000). In Asian isolates, the fourth box is not as evident as in the isolates from America and Australia (Silva-Rosales *et al.*, 2000). The EK similarities found among Mexican and Cuban PRSV isolates support a possible relationship of these blocks with the geographical distribution of PRSV.

In phylodynamic studies, 141 PRSV CP gene sequences from continental America and the Caribbean islands available in GenBank or generated in this work were

analysed. In previous studies, Bateson *et al.* (2002) and Gibbs *et al.* (2008) provided evidence of the Asian origin of PRSV, which was supported with a phylogeographic approach by Olarte-Castillo *et al.* (2011). The results obtained reveal dispersion of the virus among locations over time. Some dispersal events from Colombia and Mexico to Venezuela are suggested, and spatial diffusion showed other events between Brazil and Jamaica, whilst only dispersions from Brazil were detected in Argentina.

In the Bayesian coalescent analysis, an apparent structure of PRSV sequences from Cuba was found in relation to the geographical distribution of this virus, which provides strong evidence for the *in situ* evolution of viruses within individual countries (Simmons *et al.*, 2008). Moreover, that structure was not observed in all the analysed sequences, indicating dispersion of the virus among locations over time, shown by Olarte-Castillo *et al.* (2011). The phylogeographical studies revealed that viruses can be subjected to dispersion within localities or between different regions, perhaps mediated by the local spread of aphids in nearby areas and the movement of infected papaya seedlings, which can spread the virus over long distances. However, isolates sampled from adjoining locations were not always related, as observed among some Cuban isolates, suggesting that biogeographical structure may be determined by the trading of infected papaya seedlings, which has been reported for other viruses (Desbiez *et al.*, 2002). Accordingly, Lecoq *et al.* (2003) demonstrated the potential of cucurbit fruit as potyvirus transmission sources over long distances. In Cuba, PRSV-P infecting cucurbits has not been detected; however, cucurbits can also be a potyvirus transmission source and were reported as potential hosts of this virus biotype (Mansilla *et al.*, 2012). On the other hand, surveys of producers and seedling production nurseries revealed movement of genetic resources between provinces, which supports the observations with the Palma Soriano (MF041962) isolate.

According to the results obtained in this work regarding the demographic reconstruction, the increment of viral diversity observed from 1985 to 1990 coincided with periods of implementation of extensive production practices in the region. Based on FAO data, the increase in the harvested papaya area was notable between 1980 and 1990, which approximately coincided with reports of severe epidemics of PRSV in Jamaica, Cuba, Venezuela, Mexico, Puerto Rico and Brazil.

Phylodynamic analysis provides a reliable estimate of phylogeny and divergence times of PRSV under uncorrelated relaxed-clock models (Drummond *et al.*, 2006). Olarte-Castillo *et al.* (2011) reported an exponential increase of PRSV diversity over time. Nevertheless, the present results with isolates from continental America and the Caribbean islands showed that the viral diversity of this population showed a slight increase from 1944 to 1951, remained almost constant until 1985, and then increased again for a short period, before remaining constant for the last 25 years. Olarte-Castillo *et al.* (2011) dated the origin of PRSV diversification in the Americas at between 250 and 100 years ago. However, in the present analysis using some of the previously evaluated sequences and new isolates from continental American (four from Argentina) and the Caribbean (34 from Cuba, 27 from Puerto Rico and 12 from Jamaica), the estimates were approximately 100–75 years ago (between 1916 and 1941). This difference may be attributed not only to the differences in the dataset, but also to the different methods applied in both studies. In this work, the flexible molecular clock and demographic models used allowed complex evolutionary scenarios to be evaluated.

The evidence presented of the epidemic growth estimation by means of demographic reconstruction analysis of PRSV provides important information for the implementation of control strategies based on the use of transgenic plants. These results suggest that practices such as the use of transgenic plants could be influenced by the dynamics of the viral population (Fermin *et al.*, 2010). This encourages the application of phylodynamic studies to monitor the epidemiological patterns. Finally, the geographical structure observed among Cuban isolates, which were grouped into eastern and western-central regions, suggests the establishment and conservation of genetic variants into a particular population. The PRSV distribution in the main papaya-growing areas in Cuba reveals the importance of developing effective virus disease management programmes that include nursery protection, crop rotation, crop barriers, management of host plants and vectors signalling to mitigate the high virus incidence in these important papaya-producing regions. According to Fargette *et al.* (2008), any evolution, as might occur in a species jump, should be accounted for by a relaxed-clock model. The molecular dating technique provides insights into the history of viruses; therefore, in this study phylodynamic reconstruction and spatiotemporal diffusion processes of PRSV were incorporated as an important epidemiological tool for PRSV management.

Acknowledgements

This study was supported in part by the International Foundation for Science, Stockholm, Sweden, through a grant to D.C.M. (D/5134-1). The authors want to thank Yulexis Pino and Danilo Díaz (AGROFAR, Cuba) for their support during the collection of papaya leaves across the country. The authors have no conflict of interest to declare.

References

Acuña J, Zayas F, 1946. *El Mosaico y Otras Enfermedades de la Fruta Bomba* (Carica papaya L.). *Circular 85*. La Habana, Cuba: Estación Experimental Agronómica de Santiago de las Vegas.

Ayres DL, Darling A, Zwickl DJ *et al.*, 2012. BEAGLE: an application programming interface and high-performance computing library for statistical phylogenetics. *Systematic Biology* **61**, 170–3.

Baele G, Lemey P, Bedford T, Rambaut A, Suchard MA, Alekseyenko AV, 2012. Improving the accuracy of demographic and molecular

clock model comparison while accommodating phylogenetic uncertainty. *Molecular Biology and Evolution* **29**, 2157–67.
Bateson MF, Lines RE, Revill P *et al.*, 2002. On the evolution and molecular epidemiology of the potyvirus *Papaya ringspot virus*. *Journal of General Virology* **83**, 2575–85.
Bielejec F, Baele G, Vrancken B, Suchard MA, Rambaut A, Lemey P, 2016. SpreaD3: interactive visualisation of spatiotemporal history and trait evolutionary processes. *Molecular Biology and Evolution* **33**, 2167–9.
Cabrera Mederos D, Dal Zotto A, Portal O, Giolitti F, 2016. First report of *Papaya ringspot virus* infecting *Carica papaya* in Argentina. *Journal of Plant Pathology* **98**, 687.
Cabrera Mederos D, Cruz M, Nome CF, Giolitti F, Portal O, 2017. Aggressiveness of Cuban *Papaya ringspot virus* isolates on *Carica papaya* L. cv. Maradol Roja under greenhouse conditions. *Journal of Plant Physiology and Pathology* **5**, 2.
Chin M, Rojas Y, Moret J, Fermin G, Tennant PF, Gonsalves D, 2007. Varying genetic diversity of *Papaya ringspot virus* isolates from two time-separated outbreaks in Jamaica and Venezuela. *Archives of Virology* **152**, 2101–6.
Cooke BM, 2006. Disease assessment and yield loss. In: Cooke BM, Gareth Jones D, Kaye B, eds. *The Epidemiology of Plant Diseases*, 2nd edn. Dordrecht, Netherlands: Springer, 43–80.
Crane JH, Dorey AJ, Schaffer BA, McMillan RT, 1995. Comparison of *Papaya ringspot virus* effects on 23 cultivars and 18 selections of papaya (*Carica papaya*) in South Florida. *Proceedings of the Florida State Horticultural Society* **108**, 354–7.
Darriba D, Taboada GL, Doallo R, Posada D, 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* **9**, 772.
Davis M, White TL, Crane JH, 2004. Resistance to papaya ringspot virus in transgenic papaya breeding lines. *Proceedings of the Florida State Horticultural Society* **117**, 241–5.
Desbiez C, Wipf-Scheibel C, Lecoq H, 2002. Biological and serological variability, evolution and molecular epidemiology of *Zucchini yellow mosaic virus* (ZYMV, Potyvirus) with special reference to the Caribbean islands. *Virus Research* **85**, 5–16.
Drummond AJ, Rambaut A, Shapiro B, Pybus OG, 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution* **22**, 1185–92.
Drummond AJ, Ho SY, Phillips MJ, Rambaut A, 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biology* **4**, e88.
Drummond AJ, Suchard MA, Xie D, Rambaut A, 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* **29**, 1969–73.
FAO, 2017. FAO food and agriculture data. [http://www.fao.org/faostat/]. Accessed 4 September 2018.
Fargette D, Pinel-Galzi A, Sérémé D *et al.*, 2008. Diversification of *Rice yellow mottle virus* and related viruses spans the history of agriculture from the neolithic to the present. *PLoS Pathogens* **4**, e1000125.
Fermin G, Castro LT, Tennant PF, 2010. CP-transgenic and non-transgenic approaches for the control of papaya ringspot: current situation and challenges. *Transgenic Plant Journal* **4**, 1–15.
Fernández-Rodríguez T, Rubio L, Carballo O, Marys E, 2008. Genetic variation of papaya ringspot virus in Venezuela. *Archives of Virology* **153**, 343–9.
Gibbs AJ, Ohshima K, Phillips MJ, Gibbs MJ, 2008. The prehistory of potyviruses: their initial radiation was during the dawn of agriculture. *PLoS ONE* **3**, e2523.
Holman J, 1974. *Los Áfidos de Cuba*. La Habana, Cuba: Instituto Cubano del Libro.
Jain RK, Sharma J, Sivakumar AS *et al.*, 2004. Variability in the coat protein gene of *Papaya ringspot virus* isolates from multiple locations in India. *Archives of Virology* **149**, 2435–42.
Kumar S, Stecher G, Tamura K, 2016. mega 7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution* **33**, 1870–4.
Lecoq H, Desbiez C, Wipf-Scheibel C, Girard M, 2003. Potential involvement of melon fruit in the long distance dissemination of cucurbit potyviruses. *Plant Disease* **87**, 955–9.
Lemey P, Rambaut A, Drummond AJ, Suchard MA, 2009. Bayesian phylogeography finds its roots. *PLoS Computational Biology* **5**, e1000520.
Mansilla PJ, Moreira AG, Mello APOA *et al.*, 2012. Importance of cucurbits in the epidemiology of *Papaya ringspot virus* type P. *Plant Pathology* **62**, 571–7.
Miller MA, Pfeiffer W, Schwartz T, 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. *Gateway Computing Environments Workshop GCE 2010*. Piscataway, NJ, USA: IEEE, 1–8. https://doi.org/10.1109/gce.2010.5676129.
Noa-Carrazana JC, González-de-Leon D, Ruiz-Castro BS, Piñero D, Silva-Rosales L, 2006. Distribution of *Papaya ringspot virus* and *Papaya mosaic virus* in papaya plants (*Carica papaya*) in Mexico. *Plant Disease* **90**, 1004–11.
Olarte-Castillo XA, Fermin G, Tabima J *et al.*, 2011. Phylogeography and molecular epidemiology of *Papaya ringspot virus*. *Virus Research* **159**, 132–40.
Pagán I, García-Arenal F, 2018. Population genomics of plant viruses. In: *Population Genomics*. Cham, Switzerland: Springer, 1–33. https://doi.org/10.1007/13836_2018_15.
Purcifull DE, Edwardson JR, Hiebert E, Gonsalves D, 1984. *Papaya ringspot virus*. *CMI/AAB Descriptions of Plant Viruses no. 292*. Wallingford, UK: CAB International.
Rambaut A, Suchard MA, Xie D, Drummond AJ, 2014. tracer v. 1.6. [http://tree.bio.ed.ac.uk/software/tracer/]. Accessed 4 September 2018.
Rambaut A, Lam TT, Carvalho LM, Pybus OG, 2016. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evolution* **2**, vew007.
Rieux A, Khatchikian CE, 2017. TipDatingBeast: an R package to assist the implementation of phylogenetic tip-dating tests using BEAST. *Molecular Ecology Resources* **17**, 608–13.
Rodríguez D, Geraldino PdS, González J, dos Reis A, 2014. Molecular and biological studies of *Papaya ringspot virus* isolates from Brazil and Cuba. *American Journal of Agriculture and Forestry* **2**, 209–18.
Roig JT, 1965. *Diccionario Botánico de Nombres Vulgares Cubanos*, vol. II. La Habana, Cuba: Editora del Consejo Nacional de Universidades.
Ronquist F, Teslenko M, van der Mark P *et al.*, 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* **61**, 539–42.
Silva-Rosales L, Becerra-Leor N, Ruíz-Castro S, Téliz-Ortiz D, Noa-Carrazana JC, 2000. Coat protein sequence comparisons of three Mexican isolates of papaya ringspot virus with other geographical isolates reveal a close relationship to American and Australian isolates. *Archives of Virology* **145**, 835–43.
Simmons HE, Holmes EC, Stephenson AG, 2008. Rapid evolutionary dynamics of *Zucchini yellow mosaic virus*. *Journal of General Virology* **89**, 1081–5.
Townsend GR, Heuberger JW, 1943. Methods for estimating losses caused by diseases in fungicide experiments. *Plant Disease Reporter* **27**, 340–3.
Tripathi S, Suzuki JY, Ferreira SA, Gonsalves D, 2008. *Papaya ringspot virus*-P: characteristics, pathogenicity, sequence variability and control. *Molecular Plant Pathology* **9**, 269–80.
Wylie SJ, Adams M, Chalam C *et al.*, 2017. ICTV Virus Taxonomy Profile: *Potyviridae*. *Journal of General Virology* **98**, 352–4.
Xia X, 2017. dambe6: new tools for microbial genomics, phylogenetics and molecular evolution. *Journal of Heredity* **108**, 431–7.
Xia X, Xie Z, Salemi M, Chen L, Wang Y, 2003. An index of substitution saturation and its application. *Molecular Phylogenetics and Evolution* **26**, 1–7.
Yasaka R, Fukagawa H, Ikematsu M *et al.*, 2017. The time of emergence and spread of turnip mosaic potyvirus. *Scientific Reports* **7**, 4240.

Yeh SD, Bau HJ, Kung YJ, Yu TA, 2007. Papaya. In: Pua EC, Davey MR, eds. *Biotechnology in Agriculture and Forestry*, Vol. 60. Transgenic Crops V. Berlin, Germany: Springer-Verlag, 73–96.

Zambrana-Echevarría C, De Jesús-Kim L, Márquez-Karry R, Siritunga D, Jenkins D, 2016. Diversity of *Papaya ringspot virus* isolates in Puerto Rico. *HortScience* **51**, 362–9.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site.

Figure S1. Analysis of the date-randomization test on the time to the most recent common ancestor parameter.

Figure S2. Analysis of the date-randomization test on the (mean) substitution rate parameter.

Figure S3. Papaya ringspot virus symptoms induced in papaya plants cv. Maradol Roja cultivated under field conditions in Cuba. On leaves: (a) vein clearing, (b) mild mosaic, (c) patches, (d) mottling, (e) deformation, (f) dark green swelling zones, (g) intense mosaic, (h) leaf distortion. On stems and petioles: (i) oily spots. On fruits: (j): concentric rings and deformation.

Figure S4. Geographical distribution of papaya ringspot virus on papaya plants cv. Maradol Roja in Cuba. Locations where viral disease was detected are numbered from 1 to 36 (shown in Table 2). Provinces of interest are labelled. PR: Pinar del Río, AR: Artemisa, LH: La Habana, MY: Mayabeque, MT: Matanzas, CF: Cienfuegos, VC: Villa Clara, SS: Sancti Spíritus, CM: Camagüey, LT: Las Tunas, HG: Holguín, GR: Granma, SC: Santiago de Cuba.

Figure S5. Pairwise analysis among nucleotide and amino acid sequences of the coat protein gene of papaya ringspot virus isolates from Cuba. Multiple alignments of sequences were performed with MEGA 7 software.

Figure S6. The glutamic acid and lysine (EK) repeat patterns (I, II, III, IV) in aligned sequences of the papaya ringspot virus (PRSV) coat protein from Cuba and Mexico isolates. Shaded boxes indicate EK repeat residues.

Table S1. Plantations of papaya cv. Maradol Roja assessed in Cuba.

Table S2. Accession numbers of coat protein gene sequences of papaya ringspot virus (PRSV-P) used in this study.

Table S3. Bayes factors (BF) for comparison of molecular clocks and coalescent models.

Table S4. Supported transition rates among locations obtained from the analysis of the spatiotemporal dispersion (BSSVS).